J Fam Viol (2008) 23:69–80
DOI 10.1007/s10896-007-9128-5

ORIGINAL ARTICLE

Assessing the Risk of Future Psychological Abuse: Predicting the Accuracy of Battered Women's Predictions

Margret E. Bell · Lauren Bennett Cattaneo · Lisa A. Goodman · Mary Ann Dutton

Published online: 4 October 2007

Abstract Building on a handful of studies demonstrating battered women's accuracy in assessing their risk of being physically reabused, this study examined how accurately victims assess their risk of future psychological abuse. Participants' ratings of the likelihood that their partner would engage in controlling/dominance behaviors or efforts to humiliate/degrade them in the coming year and their reports 18 months later of whether this had actually occurred were used to create a four category version of accuracy (true positive, false positive, true negative, false negative). Victims were more likely to be right than wrong in their assessments of risk; PTSD symptoms, the recency of physical violence, and the degree of stalking and psychological abuse in the relationship predicted membership in the four accuracy categories. These findings overlap considerably with those examining victim accuracy in predicting physical abuse and inform ongoing debates about the value of incorporating victims' insights into risk assessment efforts.

Keywords Intimate partner violence · Battered women · Psychological abuse · Risk assessment · Victim accuracy

A major task facing clinicians, researchers, and policy-makers working in the area of intimate partner violence (IPV) is to learn ways to more effectively balance our knowledge and expertise with victims' knowledge and expertise. Far too often the field has gone to extremes—in the past, for example, failing to assist women seeking legal sanctions against their partner and more recently, implementing mandatory arrest and prosecution policies that take important choices out of victims' hands—rather than partnering and working collaboratively with women.

For many, a serious obstacle to these types of collaborative partnerships is concern about the degree to which victims are able to realistically appraise their risk of experiencing reabuse. Specifically, practitioners often struggle with how much weight a victim's perceptions of her risk should be given in estimating the level of danger and/or a developing a plan for intervention, particularly relative to professional assessments (Bennett Cattaneo and Goodman 2007).

Although only a handful of empirical studies have addressed this issue, those that have found that victims were at least as, if not more, accurate in predicting their risk of being physically reassaulted than were practitioners, risk factors identified by prior research, and all but one of the standardized risk assessment instruments investigated (Bennett Cattaneo 2007; Bennett Cattaneo and Goodman 2003; Heckert and Gondolf 2004; Roehl et al. 2005; Weisz et al.

M. E. Bell (✉)
Military Sexual Trauma Support Team, Department of Veterans Affairs Office of Mental Health Services and National Center for PTSD, VA Boston Healthcare System,
150 S. Huntington Ave., 116B-3,
Boston, MA 02130, USA
e-mail: mebell@aya.yale.edu

L. B. Cattaneo
Department of Psychology, George Mason University,
Fairfax, VA, USA

L. A. Goodman
Department of Counseling, Developmental, and Educational Psychology, Lynch School of Education, Boston College,
Chestnut Hill, MA, USA

M. A. Dutton
Department of Psychiatry, Georgetown University,
Washington, DC, USA

2000). This body of work has also shown that including victim assessments improves predictions of future violence above and beyond the accuracy attained by using empirically identified risk factors or risk assessment instruments.

Most recently, the authors of the current study compared victims' predictions about their risk of physical reabuse with the abuse they actually experienced over the course of 18 months (Bennett Cattaneo et al. 2007). In contrast to previous studies which have used dichotomous operationalizations of accuracy (i.e., victims are either right or wrong) or receiver operating characteristic (ROC) analysis (which though state of the art in other ways, ultimately yields a single number indicating degree of accuracy), we used a four-category version of accuracy (true positive, true negative, false positive, and false negative). This decision was based on the assumption that there would be qualitative differences between ways of being correct (predicting high risk and being reabused vs predicting low risk and not being reabused) and incorrect (predicting high risk and not being reabused vs predicting low risk and being reabused). Our findings revealed that overall, women were more likely to be accurate than they were to be inaccurate. They were not biased towards being accurate or inaccurate in certain ways, however—that is, they did not tend to be overly optimistic or pessimistic. In addition, we identified several variables that predicted membership in the four accuracy categories: history of physical violence and stalking from the current partner, physical violence from former partners, current level of substance use, current posttraumatic stress disorder symptoms (PTSD) and the recency of the violence. Among the most policy-relevant findings were the increased likelihood that the victim would be a "true positive" if she had been more frequently stalked by her partner and that she would be a "false negative" if she reported higher levels of substance use.

Though the few existing studies reviewed above have made crucial contributions to our knowledge base in this area, they have all focused exclusively on physical abuse. To our knowledge, no studies have examined battered women's accuracy in assessing their risk of experiencing future psychological abuse. Why might this information be important? First, one key component to fostering collaborative partnerships with victims is an increased focus on issues that victims themselves define as important. For example, though battered women report that they find their partner's psychologically abusive behavior to be more damaging to their emotional well-being than their experiences of physical violence (Follingstad et al. 1990), psychological abuse continues to be generally understudied relative to physical abuse. This is despite the well-established link between psychological abuse and later physical abuse (Kelly 2004; Stith et al. 2004) and between psychological abuse and numerous deleterious outcomes, including PTSD, depression, drug use, low self-esteem, and poor physical health, even after accounting for the effects of physical abuse (Arias 1999; Arias and Pape 1999; Katz and Arias 1999; Kelly 2004).

Second, the presence and severity of psychological abuse is in some ways a core indication of the overall quality of a victim's relationship with her partner. Psychologically abusive behaviors often occur on a much more frequent basis than do acts of physical violence and indeed, may permeate interactions between victim and batterer on a daily basis (Marshall 1999; Pence and Paymar 1993). Given oft-raised concerns about victims being naive about the long-term prognosis for the abuse in their relationship, it seems particularly important to understand how accurately battered women are able to assess the likelihood of this form of mistreatment continuing over time.

Finally, studies exploring the connection between physical abuse and psychological abuse have shown that though the two are often considerably intertwined, they also exhibit dynamics and patterns that do not appear to be entirely overapping (Jacobson et al. 1996; O'Leary 1999). It thus remains unclear whether findings regarding victims' ability to accurately predict physical abuse will generalize to the prediction of psychological abuse. Understanding more about the latter is clearly important in its own right, for the reasons outlined above. However, the strong relationship between the two forms of abuse also indicates the potential utility of drawing on information related to both types of risk assessments when designing research and developing intervention strategies. In the current study, which we describe shortly, we draw on both of these approaches in order to add to our knowledge about how these two forms of abuse are both related and distinct.

It is important to note that although psychological abuse has been defined in different ways and noted to manifest itself in many interrelated forms (e.g., verbal harassment/criticism; withheld or contingent emotional support; extreme jealousy; threats of harm to self or the victim; economic deprivation), controlling/dominance and humiliation/degradation have been identified as particularly toxic for victims' physical and mental health; they are also among the most frequently experienced forms of psychological abuse and the forms most strongly associated with physical violence (Bennett et al. 2000; Follingstad et al. 1990; Katz and Arias 1999; Kelly 2004; Murphy and Hoover 1999). As such, we focus this article on these two types of psychological abuse in particular.

Research Questions and Predictors

In an effort to address the issues identified above, the current study sought to extend our previous study by addressing the

same two questions—(1) how accurate are victims in predicting reabuse? and (2) what predicts victim accuracy in predicting reabuse?—in relation to psychological abuse. Although the primary emphasis of this paper is on psychological abuse, our discussion of this study's findings will also draw on the results from our previous paper on physical abuse in order to consider the degree to which there is overlap between the two types of accuracy. This seems important as in practice, professionals are likely to be more interested in the general risk posed by the batterer than the risk that he will engage in a specific behavior.

In service of this goal, we chose to utilize the same predictors in addressing our second question as we used in our previous study, in order to allow comparisons between accuracy in predicting physical reabuse and accuracy in predicting psychological reabuse. A more detailed rationale for the selection of these particular predictors is outlined in our previous paper, but in brief, given the absence of a relevant overarching theory, we used Bronfenbrenner's (1988) ecological framework to select factors at the individual, interpersonal, and systemic levels that prior research and/or logic suggest are likely to influence accuracy. We view this study as exploratory and thus refrain from making specific hypotheses about the connection between predictors and particular categories of accuracy. That being said, it seems likely that variables at the individual level that affect a victim's mental state, including *symptoms of PTSD*, *substance use*, and the *closeness in time to the most recent incident of physical violence*, might also affect the clarity with which she views her situation. To the extent that risk assessment is a skill that can be learned, it will likely be affected by interpersonal-level variables such as the *duration of the victim's relationship* with the abuser, the *history of abuse in the relationship*, and her *experiences of abuse in other relationships*. Also at the interpersonal level, both *available social support*, or the degree to which a social support network exists, and the actual *use of the support network* are likely important in helping victims have access to perspectives that counter the batterer's version of reality. Contact with systemic-level sources of social support, as obtained through *formal help-seeking* from the police, mental health professionals, shelters, or members of the clergy, should operate similarly.

Method

Procedure

This study draws on data from the first interview (Time 1) and 18 month follow-up (Time 6) of a four-and-a-half year longitudinal study of 406 female IPV victims seeking help from shelter, civil court, and/or criminal court in a mid-Atlantic city. In that larger study, participants were recruited during the first 30 days of their stay in shelter (n=68); as they began the process of petitioning for a civil protection order (n=220); or as they exited the courtroom following the final disposition of a criminal case against their abusive partner (n=118). Eligible participants were over 18, English-speaking, and without significantly impaired mental status at first contact with interviewers. Of those approached, 116 (29%) women refused to participate, largely due to time constraints; this differed significantly across sites, with women approached at criminal court being less likely and women in shelter being more likely to participate than expected. Two hundred seventy-three participants had complete data at the two time points of interest for this study (T1 and T6).

Measures

Demographic Information At Time 1, participants were asked to report general demographic information including their age, level of education, employment status, income level, and number of children. We include these data here for descriptive purposes.

Individual-Level Predictors We used the PTSD Checklist (PCLS, Blanchard et al. 1996; Weathers et al. 1993) to measure symptoms of *PTSD* at Time 1. This 17-item scale uses a one (not at all) to five (extremely) Likert scale to assess the extent to which participants have experienced PTSD symptomatology in the past month. Summing responses to produce an index of severity of symptoms, we obtained a Cronbach alpha of 0.94 for this study.

We measured *substance use* at Time 1 with three items that asked the participant to indicate how often she had "[drunk] any alcohol", had "[drunk] until she was intoxicated", or had "[used] street drugs" in the past month; response choices for these items were on a Likert scale from one (not at all) to five (almost daily). A participant's score consisted of the sum of her responses to the individual items. The Cronbach alpha was 0.68.

At Time 1, in order to measure the closeness in time to the most recent incident of physical violence (*recency of the assault*), participants were asked to report "the most recent time he did any of these [physically abusive or stalking-related] things to you." Response choices were "within the last week", "last month", "last 3 months", "last 6 months", "last 12 months", or "more than a year ago", providing an ordinal measure of recency.

Interpersonal-Level Predictors To examine the history of abuse in the victim's relationship with the current perpetrator, we asked victims to complete measures assessing experiences of psychological abuse, physical abuse, and stalking. Specif-

ically, we assessed *psychological abuse* at Time 1 with a short form of the Psychological Maltreatment of Women Inventory (PMWI-Short Form, Tolman 1989; Tolman 1999). This measure asks participants the degree to which they have experienced various behaviors representing dominance/isolation or emotional/verbal abuse in the past year. Participants' responses on a one (never) to five (a lot) Likert scale were summed to provide an overall measure of psychological abuse. The Cronbach alpha for this scale was 0.90.

To measure *physical abuse* from the current perpetrator in the year prior to Time 1, we used a yes/no version of the Conflict Tactics Scale-2 (CTS-2 Form A, Straus et al. 1996) to ask participants whether they had experienced specific acts of physical abuse, sexual abuse, or injury related to abuse in the last year. For this study, we modified the sexual abuse subscale by removing items that asked about the perpetrator "insisting" on sex without the use of force or threats and by adding a new item: "I had sex with him because I was afraid of what he would do if I didn't." These changes reduced this subscale from seven to four items. We obtained our measure of physical abuse by summing the percentage of items endorsed on each of the subscales. The Cronbach alpha for this scale was 0.88.

We modified the National Violence Against Women Survey's measure of *stalking* (VAWS, Tjaden and Thoennes 2000) by simplifying the wording of several questions, adding one item ("He hurt or killed my pet"), and eliminating two others ("He left unwanted items for me to find" and "He tried to communicate in other ways against my will"); this left us with a total of seven items. Participants responded "yes" or "no" to having experienced these forms of stalking from the current perpetrator in the year prior to Time 1; the percentage of items they endorsed served as our measure. The Cronbach alpha for this scale was 0.83.

The *length of the participant's relationship with her partner* and the *duration of the abuse* (the length of time between the first and most recent incident of physical abuse or stalking) were assessed with single questions at Time 1.

To assess her *experiences of IPV with previous partners*, each participant was asked at Time 1 whether any intimate partner other than the current perpetrator had "ever used physical violence toward" her, had "ever sexually abused" her, had "ever physically injured" her, or had ever physically injured her to the extent that she "received medical attention for injuries." Using percentage of items endorsed as our measure, the Cronbach alpha was 0.79.

At Time 1, we assessed participants' perceptions of *available social support* using the 40-item true–false Interpersonal Support Evaluation List (ISEL, Cohen et al. 1985). The ISEL consists of four subscales measuring perceived availability of four types of support: tangible, appraisal, esteem and belonging. Given the high correlations among the subscales, we followed House and Kahn's (1985) recommendation to use a total summary score. With percent endorsed of all ISEL items as our measure we obtained a Cronbach alpha of 0.93.

We assessed *use of informal support* (support from family and friends) at Time 1 with the Intimate Partner Violence Strategies Index (IPVSI), a measure developed in a prior study using this same dataset (Goodman et al. 2003). The IPVSI consists of six subscales that cover the range of strategies that victims use to respond to IPV. The informal help-seeking subscale asks whether a participant has ever used any of three strategies to cope with the violence in her relationship: talked to family and friends about what to do to protect herself or her children; stayed with family or friends; or sent her children to stay with family or friends. We used percent of items endorsed as our predictor. Though the Cronbach alpha for this subscale (0.53) suggested only moderate internal reliability, the items on this scale represent distinct behaviors under the same conceptual umbrella and do not necessarily tap into an underlying construct. This makes internal consistency an inappropriate measure of its reliability (see Goodman et al. 2003). We therefore did not view this low Cronbach alpha value as a reason to exclude the scale.

System-Level Predictors To assess *formal help-seeking* at Time 1 we combined the legal help-seeking and formal help-seeking subscales of the Intimate Partner Violence Strategies Index. This created a measure of the degree to which participants had ever sought help from the civil or criminal court system, legal aid, their workplace, clergy, medical professionals, counselors, or shelters. Percentage of items endorsed served as our predictor. The Cronbach alpha for these combined subscales was 0.84.

Dependent Variable: Victim Accuracy Assigning women to our four accuracy categories required several steps and some difficult decisions. Two overarching goals shaped the final outcome: first, we were committed to constructing the four categories (true positive, etc.) described earlier given our belief that previous operationalizations have missed important nuances in this variable. Second, as mentioned earlier, we believe that in practice, professionals are likely to be interested in a victim's assessment of the general risk posed by the batterer rather than the risk that he will engage in a specific behavior. We therefore collapsed across items both in the victim's assessment of risk and in any reabuse she experienced.

More specifically, to create our accuracy categories, we first used the Time 1 data to determine each participant's overall assessment of her *risk of being reabused* during the next year. Victims rated on a one (low risk) to five (high risk) scale the likelihood that their partner would (1) 'control or dominate'; or (2) 'humiliate or degrade' them in the next year. We combined ratings of risk on these two items (r=0.80)

into a sum score ranging from two to ten. We then divided participants into groups based on whether they predicted a lower (five and below) or higher (seven and above) risk of violence. We dropped from consideration those participants whose assessments were at the midpoint (6; n=29), rather than arbitrarily placing them in one category or another[1].

Next, we examined the *reabuse actually experienced* by each participant during the next 18 months, as reported at Time 6[2]. Again, we were interested in the victim's ability to predict any reabuse, as opposed to reabuse of a specific type. We thus chose to collapse across our two psychological reabuse items (φ=0.58) to create a single dichotomous item (did/did not experience at least one type of psychological reabuse).

Finally, using victims' Time 1 assessments of risk and their reabuse status at Time 6, we divided the 244 women in our sample into the four accuracy categories (true/false positive, true/false negative) summarized in Table 1.

Table 1 Summary of accuracy categories

	Reabused before Time 6		
Time 1 prediction of risk	No	Yes	Total
Low	True negatives 85 (69.7 expected)	False negatives 49 (64.3 expected)	134
High	False positives 42 (57.3 expected)	True positives 68 (52.7 expected)	110
Total	127	117	244

Cell values = number of participants in that category

Results

Description of Participants

Demographics and Predictors More than half (59%) of the women in our sample were recruited as they sought a civil protection order against their partners; 28% and 13% were recruited from criminal court and shelter, respectively. Participants were overwhelmingly Black (80%) and deeply impoverished: despite the fact that 60% were employed either full-or part-time, 92% had a personal income of less than $30,000 a year; 67% made less than $15,000 a year. Increasing their economic burden, 89% were raising at least one child. On average, participants were 33 years old (SD= 8.48), and almost three-quarters (74%) had completed high school. With regard to relationship status at Time 1, 71% of women were still together with their partner (40% dating; 31% married) and 29% were estranged from him (20% formerly dating him; 7% married but separated; 1% divorced).

Table 2 summarizes the characteristics of the sample with respect to our predictor variables. It is worth highlighting that levels of abuse were high in the year prior to Time 1: all but one participant reported experiences of psychological abuse and 87% reported at least one act of severe physical violence such as being kicked, "beat up," assaulted with a knife or gun, or threatened or forced into sex.

Risk, Reabuse and Accuracy At T1, participants rated the risk that the abuser would control or dominate them in the next year an average of 2.60 out of 5 (SD=1.66) and the risk that he would try to humiliate or degrade them an average of 2.82 out of 5 (SD=1.71). When we summed these items and dichotomized the result into predictions of low versus high risk as described earlier, about half of the women in our sample (55%) felt they were at relatively low risk of being psychologically reabused; 45% felt they were at high risk. At Time 6, 34% of women stated that their abuser had controlled or dominated them in the last 18 months; 41% said he had humiliated or degraded them. Collapsing these categories, 52% of our sample had been reabused by Time 6, while 48% had not. Table 1 summarizes this information.

Generalizability of the Sample

As noted above, participants from the overarching study (N=406) were dropped from our analyses for one of two reasons. First, 133 (31%) of the original participants either did not have complete Time 1 risk data (n=8) or were not reached at the 18-month follow-up (n=125). Chi-square and ANOVA analyses revealed that these participants differed from our sample only with respect to recruitment site and

[1] It is important to note that these women were not necessarily individuals who were unsure about what would happen during the next year. Given our use of a sum score, some of these participants were actually quite sure that one event would happen in the next year (rating of 5) but also quite sure that another would not (rating of 1). Analyses not detailed here revealed that dropping participants with sum scores of 6, including them with participants whose sum scores were from 1 to 5, and including them with participants whose sum scores were from 7 to 10 all resulted in very similar findings.

[2] It would have been optimal to have a more precise match between the time frame for which Time 1 risk was assessed (one year) and the time period for which reabuse was assessed (18 months). This was not feasible given the data available. We felt however that victims' assessments of risk over the next year and their assessments of risk over the next 18 months would likely be very similar, such that the mismatch would have minimal impact on the validity of our results.

Table 2 Sample descriptives at Time 1

Variable	Mean	SD	Range in sample
PTSD symptoms	47.55	18.17	17–85
Substance use	4.39	2.19	3–15
Recency of the violence	2.20	1.35	1–6[a]
Length of relationship	76.21 months	75.86 months	1.5–427 months (6 weeks–35 1/2 years)
Duration of the abuse	23.32 months	41.37 months	0–288 months (0 months–24 years)
Psychological abuse	44.55	13.61	13–65
Physical abuse[b]	0.42	0.24	0–0.95
Stalking[b]	0.36	0.28	0–1
Previous partner violence[b]	0.28	0.34	0–1
Social support available[b]	0.69	0.23	0.1–1
Use of informal support[b]	0.56	0.33	0–1
Use of formal support[b]	0.47	0.26	0–1

[a] Ordinal measure
[b] Percent endorsed (see "Measures" section)

the length of their abusive relationship. More specifically, participants recruited from the criminal court and shelter were significantly more likely than expected to have missing data; the opposite was true for participants recruited while obtaining a restraining order [$\chi^2(2)=9.84$, $p=0.007$]. Participants with incomplete data were also marginally more likely to have been involved with their abusive partner for a shorter period of time than were participants included in our analyses [$F(1, 374)=3.44$, $p=0.07$].

Second, 29 (7%) participants with complete data were dropped because their total score was at the midpoint (6) of the risk assessment scale. Participants excluded for this reason were significantly more likely to report having endured violence in their relationship for a longer period of time [$F(1, 271)=4.92$, $p=0.03$] and marginally more likely to report that the latest violent episode had been more recent [$F(1, 271)=3.37$, $p=0.07$] than were participants included in our analyses. As noted in Footnote 5, dropping these participants had a negligible impact on our main findings, which are described next.

How Accurate Are Victims?

In the first of the study's primary analyses, we used the Fisher's Exact and MeNemar's chi-square tests on our risk prediction (low/high) and reabuse (yes/no) variables to help us determine whether participants were evenly distributed across accuracy category groups (Given that our groups were based on reabuse status, we could not directly compare the absolute numbers in our four categories). In separate analyses, these tests collapse across cells on the off-diagonal (Fisher's Exact) and diagonal (McNemar) to compare the number of observed cases with the number of cases that would be expected based on the distribution of the variables in the sample as a whole.

Using Fisher's Exact test to determine whether participants fell into the "correct" (true positive/negative) or "incorrect" (false positive/negative) categories more than we would expect based on the distribution of responses in

Table 3 Correlations among predictor variables

Predictor	1	2	3	4	5	6	7	8	9	10	11
PTSD symptoms	–										
Substance use	0.22**	–									
Recency of the violence	−0.10	0.09	–								
Length of relationship	0.10	0.05	0.09	–							
Duration of the abuse	0.16*	−0.04	0.02	0.36**	–						
Psychological abuse	0.53**	0.12	−0.17**	−0.03	0.21**	–					
Physical abuse	0.48**	0.18**	−0.05	−0.01	0.26**	0.55**	–				
Stalking	0.36**	0.21**	−0.06	0.01	0.08	0.53**	0.47**	–			
Previous partner violence	0.21**	0.04	−0.01	−0.14*	−0.06	0.17**	0.22**	0.16*	–		
Social support available	−0.35**	−0.24**	0.11	0.06	−0.11	−0.26**	−0.29**	−0.10	−0.16*	–	
Use of informal support	0.29**	0.07	−0.15*	0.04	0.06	0.34**	0.17**	0.24**	0.13*	−0.07	–
Use of formal support	0.24**	0.08	−0.02	0.01	0.03	0.32**	0.30**	0.29**	0.23**	−0.11	0.48**

*$p<0.05$
**$p<0.01$

Table 4 Univariate likelihood ratios for predictors of accuracy

Predictor	χ^2	*df*	*p* value
PTSD symptoms	40.60	3	0.0001
Substance use	4.62	3	0.20
Recency of the violence	16.59	3	0.001
Length of relationship	2.10	3	>0.25
Duration of the abuse	3.23	3	>0.25
Psychological abuse	65.49	3	0.0001
Physical abuse	25.69	3	0.0001
Stalking	35.79	3	0.0001
Previous partner violence	2.92	3	>0.25
Social support available	7.78	3	0.05
Use of informal support	29.94	3	0.0001
Use of formal support	13.82	3	0.003

the sample as a whole, we found that the chi-square was significant (p=0.0001), with victims more likely to be correct than incorrect. Results from the McNemar test, however, revealed that participants were not likely to be correct or incorrect in specific ways (i.e., true positive vs. true negative; false positive vs. false negative; p=0.53).

What Predicts Accuracy Category Membership?

We used unordered multinomial logistic regression to conduct the second of the study's primary analyses. An extension of logistic regression to cases where the dependent variable has more than two categories, multinomial logistic regression examines the ability of predictors to distinguish participants in a reference category from participants in each of the remaining categories. Since we were interested in the relationships among all of our four accuracy groups, we allowed each group to take a turn serving as the reference category.

Univariate Analyses Table 3 reports the bivariate correlations for our predictors; Table 4 presents the likelihood ratios obtained in the univariate multinomial logistic regressions. At the univariate level, all but three predictors (length of the relationship, duration of the abuse, and previous partner violence) were significantly related to accuracy category membership at the p<0.25 level and were thus included in the multivariate analyses (Hosmer and Lemeshow 1989). Notably, all but one (substance use) of the predictors significant at p<0.25 were also significant at the p<0.05 level.

Multivariate Analyses Tables 5 and 6 present the multivariate multinomial logistic regression model, which considered predictors' effects relative to each other. Overall, the model fit the data well, as indicated by the non-significant Pearson and deviance chi-squares and the significant model chi-square [Pearson $\chi^2(702)=705.81$, p=0.45; deviance $\chi^2(702)=543.34$, p=1.00; model $\chi^2(27)=114.81$, p=0.0001]. The model accounted for 40% of the variance using Nagelkerke's R^2 (Tabachnick and Fidell 2001).

Using likelihood ratio tests to compare the relative contribution of each predictor, we found that the inclusion of the PTSD, recency of the violence, psychological abuse and stalking variables each significantly improved model fit. Among these predictors with significant likelihood ratios, the Wald chi-square statistic was used to determine which accuracy groups a given predictor significantly distinguished from one another. In this regard, when all other variables were held constant, a one standard deviation increase in symptoms of *PTSD* indicated that a participant was 1.68 times more likely to be a true positive than a true negative. Participants with more PTSD symptoms were also 1.96 times more likely to be false positives than true negatives and 1.77 times more likely to be false positives than false negatives. When the latest incident of violence was less *recent*, participants were particularly unlikely to be false positives: that is, they were 1.64 times more likely to be true positives than false positives, 2.19 times more likely to be true negatives than false positives, and 2.01 times more likely to be false negatives than false positives. More previous *psychological abuse* in her relationship with the abuser indicated that a participant was 2.74 times more likely to be a true positive than a true negative, 2.44 times more likely to be a true positive than a false negative, 2.15 times more likely to be a false positive than a true negative, and 1.92 times more likely to be a false positive than a false negative. Finally, more *stalking* indicated that a participant was 2.29 times more likely to be a true positive than a false positive and 1.71 times more likely to be a false negative than a false positive.

Table 5 Multivariate likelihood ratios for predictors of accuracy

Predictor	χ^2	*df*	*p* value
PTSD symptoms	8.24	3	0.04
Substance use	1.68	3	ns
Recency of the violence	10.32	3	0.02
Psychological abuse	16.59	3	0.001
Physical abuse	2.45	3	ns
Stalking	11.08	3	0.01
Social support available	5.98	3	0.11
Use of informal support	4.11	3	ns
Use of formal support	1.68	3	ns
Overall model	114.81	27	0.0001

Table 6 Multivariate parameter estimates, standard errors, Wald values, and odds ratios for prediction of accuracy

Predictor	Comparison					
	True negative vs false negative	True negative vs false positive	True negative vs true positive	False negative vs false positive	False negative vs true positive	False positive vs true positive
PTSD symptoms	−0.001 (0.01); 0.18; 0.90	−0.003 (0.02); 5.93**; 0.51	−0.03 (0.01); 4.43*; 0.59	−0.003 (0.02); 3.79*; 0.57	−0.002 (0.01); 2.57; 0.66	0.0008 (0.02); 0.33; 1.17
Substance use	−0.0001 (0.10); 0; 1.00	0.15 (0.13); 1.32; 1.39	0.03 (0.09); 0.09; 1.07	0.15 (0.13); 1.29; 1.38	0.003 (0.09); 0.09; 1.06	−0.12 (0.12); 1.01; 0.77
Recency of the violence	0.01 (0.14); 0.21; 1.09	0.58 (0.20); 8.30**; 2.19	0.22 (0.15); 1.99; 1.34	0.52 (0.21); 6.08**; 2.01	0.15 (0.16); 0.92; 1.22	−0.37 (0.21); 3.14; 0.61
Psychological abuse	−0.001 (0.02); 0.20; 0.89	−0.005 (0.02); 6.50**; 0.46	−0.07 (0.02); 12.10**; 0.36	−0.004 (0.02) 3.98*; 0.52	−0.006 (0.02); 8.24**; 0.41	−0.002 (0.02); 0.56; 0.79
Physical abuse	−1.18 (1.08); 1.21; 0.75	−1.25 (1.16); 1.17; 0.74	−0.05 (1.08); 0.003; 0.99	−0.006 (1.22); 0.003; 0.98	1.13 (1.13); 1.00; 1.31	1.19 (1.11); 1.15; 1.33
Stalking	−0.30 (0.86); 0.12; 0.92	1.62 (0.98); 2.69; 1.57	−1.35 (0.86); 2.45; 0.69	1.91 (1.02); 3.49; 1.71	−1.05 (0.89); 1.40; 0.75	−2.96 (0.93); 10.07**; 0.44
Social support available	0.88 (0.91); 0.94; 1.22	−1.82 (1.11); 2.69; 0.66	−0.21 (0.93) 0.06; 0.95	−2.71 (1.14); 5.64*; 0.54	−1.10 (0.93); 1.39; 0.78	1.61 (1.06); 2.29; (1.45)
Use of informal support	−0.42 (0.66); 0.40; 0.87	−1.34 (0.78); 2.96; 0.64	−1.18 (0.72); 2.73; 0.68	−0.92 (0.83); 1.24; 0.74	−0.77 (0.75); 1.04; 0.78	0.16 (0.80); 0.04; 1.05
Use of formal support	−0.76 (0.88); 0.74; 0.82	0.51 (0.98); 0.27; 1.14	−0.11 (0.9); 0.01; 0.97	1.27 (1.02); 1.56; 1.39	0.65 (0.92); 0.50; 1.18	−0.62 (0.92); 0.45; 0.85

Cell values = beta (standard error); Wald; odds ratio. Odds ratios are calculated in reference to the second group listed in each column
*$p<0.05$
**$p<0.01$

Discussion

Review of Findings

This exploratory study used a four category version of accuracy (true positive, false positive, false negative, true negative) to examine two questions: (1) how accurate are victims in predicting their risk of experiencing psychological reabuse?; and (2) what predicts victim accuracy in predicting psychological reabuse? In the "Implications for Practice and Policy" section of this Discussion, we also draw on the results from our previous paper on women's accuracy in predicting physical abuse in order to consider the degree to which there is overlap between the two types of accuracy.

Rates of Accuracy With regard to this first question, the findings of the current study showed that almost two thirds (62%) of victims accurately assessed their risk of being psychologically reabused, in that victims were disproportionately likely to fall into the two "correct" categories (true positives and true negatives) as compared to the "incorrect" categories (false positives and false negatives). Though the research reviewed earlier defined and measured accuracy in varying ways, it is notable that the rate of accuracy in predicting psychological abuse reported here is similar to that reported in studies assessing victims' accuracy in predicting risk of physical reabuse (63%, 74%, 64%, and 66% in Campbell 1995; Bennett Cattaneo and Goodman 2003; Heckert and Gondolf 2004; and Bennett Cattaneo et al. 2007, respectively). Additionally, we found that victims were not likely to be correct or incorrect in specific ways - that is, they were as equally skilled in predicting reabuse as they were in predicting no reabuse; and if wrong, were equally likely to overestimate as to underestimate their risk. Overall then, this study provides further evidence to suggest that battered women are more likely to be right than wrong in their appraisals of the likelihood that they will continue to experience abusive behavior from their partner. As such, their risk assessments may prove a useful source of information for clinicians and others partnering with them to ensure their safety and well-being.

Predictors of Accuracy Clearly, however, women varied in terms of how accurately they assessed their risk of reabuse. At the univariate level, most of the predictors of accuracy examined in this study were significant, suggesting that they have some bearing on accuracy and are thus useful for practitioners to assess in their work with victims. However, when the impact of these variables was examined relative to each other in a multivariate model, the recency of the latest episode of physical abuse, PTSD symptomatology, and the

history of psychological abuse and stalking in the relationship emerged as most predictive of membership in the four accuracy categories.

With regard to these variables significant in the multivariate model, the latest episode of physical violence having been more remote was associated with victims being particularly unlikely to overestimate their risk of experiencing future psychological abuse—that is, they were significantly more likely to be true negatives, false negatives, and true positives than to be false positives. They were not, however, likely to underestimate their future risk of abuse, as evidenced by the nonsignificant comparison between the false negatives and true positives for this predictor. It thus seems that while the passage of time does not have an impact on the accuracy of victims' predictions, it is associated with the type of error victims are likely to make if incorrect: when wrong, victims with more remote histories of violence are more likely to underestimate than overestimate their risk.

Also at the individual level in ecological framework terms, when all other variables were accounted for, women with more symptoms of PTSD were significantly more likely to estimate their risk of future psychological abuse to be high than to be low. Specifically, higher levels of PTSD were associated with victims being more likely to be both true positives and false positives than true negatives. Interestingly, however, women with more PTSD symptoms were correct in their predictions of high risk about as often as they were wrong in these predictions (as indicated by the nonsignificant contrast between false positives and true positives for this predictor). It seems that though one of the hallmarks of PTSD is hypervigilance and a tendency to selectively attend to, process, and remember threatening cues and information (Litz and Keane 1989), this hypervigilance does not impair victims' judgments about risk of psychological reabuse to a significant degree.

At the interpersonal level, psychological abuse exhibited a pattern similar to that noted above for PTSD. That is, women with more severe histories of psychological abuse were more likely to estimate their risk for experiencing future psychological abuse to be high, but were about equally likely to be correct (true positive) as incorrect (false positive) in those predictions.

In contrast, when women with high levels of stalking predicted their future risk of abuse to be high, they were correct more often than not (that is, they were more likely to be true positives than false positives). Experiences of stalking thus appear to provide meaningful information to victims attempting to predict their future risk of being controlled or humiliated.

Surprisingly, no variables indexing use of or access to social support or formal assistance emerged as significant in the multivariate model. This did not appear to be due to multicollinearity, as correlations between the three measures were very low (r from −0.07 to −0.11) with the exception of the relationship between use of informal support and formal help-seeking (r=0.48). Also, rerunning our analyses with only one or with various combinations of two of these variables produced findings similar to those reported here. A more likely explanation for our findings is that our operationalization of available social support was too broad in that it did not focus on assistance directly targeting a victim's abusive experiences; oppositely, the behaviors included in our measure of use of support from family and friends were quite specific and many not have tapped aspects of support impacting accuracy. Finally, all of our participants were seeking help and the vast majority were court-involved, thus perhaps limiting our ability to see the impact of system-level interventions on victim accuracy.

Limitations and Directions for Future Research on Psychological Abuse Accuracy

As this study is the first to examine victims' accuracy in predicting future experiences of psychological abuse, the generalizability of our findings beyond the particularities of our sample (low income; largely Black; help-seeking; recruited based on their experiences of physical violence) and operationalization of accuracy remains to be determined by future studies in this area. Understanding more about the reasons victims made the assessments they did may help clarify mechanisms driving the findings reported here—some victims may have made predictions based on their expectation that they would soon end their relationship with their partner; if this did not end up being the case, their inaccuracy in predicting reabuse may have been due more to this mistaken prediction than lack of skill in risk assessment. The impact that such a change in context may make in women's assessments of risk (Gondolf and Heckert 2003; Langford 1996) also highlights the importance of understanding more about how accuracy might vary over time. It also raises the question of what role victims' perceptions of the violence in their lives—their "stage of change" (Brown 1997; Prochaska and DiClemente 1982)—might play in their assessments of risk and their accuracy in prediction.

Implications for Practice and Policy

The current study and our previous paper examining battered women's accuracy in predicting physical violence are in an area of inquiry with clear implications for practice and policy. Given the newness of this topic and the exploratory nature of this study, however, we are wary of deriving too many implications from the findings reported here. In the section that follows, we therefore attempt to balance caution with efforts to lay the groundwork for

further exploration in this area and to provide useful information for practitioners. In doing so, we return to a premise we have stated repeatedly: though there are nuanced differences between women's accuracy in predicting psychological abuse and accuracy in predicting physical abuse, these two types of abuse are highly experientially and conceptually intertwined (Hall Smith et al. 2002). We thus focus this Implications section on broad inferences we can reasonably make about the accuracy of battered women's risk assessments in general, across these two forms of abuse.

So, what general implications can we draw based on our two studies of accuracy? First, more often than not, victims are correct in their assessments of risk. In both studies, women were more likely to be accurate than inaccurate in their assessments, thus suggesting that on average, victims are relatively realistic about the likelihood that their partner will continue to be abusive in a variety of forms. Moreover, victim rates of accuracy approached 66% and were comparable to those obtained by empirically derived risk factors and assessment instruments in other studies (Roehl et al. 2005), indicating that most women likely have information and insights that would contribute usefully to practitioners' risk assessment efforts. This is particularly true given that, as noted earlier, studies have shown that victim assessments add information that improves accuracy beyond what empirically-identified risk factors and risk assessment instruments are able to provide (Bennett Cattaneo and Goodman 2003; Heckert and Gondolf 2004; Roehl et al. 2005; Weisz et al. 2000). Conversely, in that about a third of women in our studies were incorrect about their level of risk, there is obviously room for improvement in many women's predictions. How professionals might best help in this regard remains unclear and is worth further study (Bennett Cattaneo and Goodman 2007).

Second, although we did not directly compare the two types of accuracy for individual women, at the aggregate level, experiences of stalking in the relationship, current symptoms of PTSD, and the recency of the latest episode of violence significantly predicted women's accuracy in making both types of assessments. Those comparisons that were significant for these variables in both studies (e.g., stalking significantly discriminated true positives from false positives in both studies) all predicted category membership in the same direction (i.e., more stalking was associated with being a true positive rather than a false positive in both studies). The overlap in these findings is not surprising given that these two forms of abuse are often so intertwined with each other in the lives of battered women. Predicting one may mean, in large part, predicting the other.

Of these variables significant in both of our multivariate models, two—symptoms of PTSD and experiences of stalking—deserve particular emphasis. With respect to PTSD, our results reveal its *lack* of connection to *inaccuracy*—that is both studies found that while more severe symptoms of PTSD were associated with predictions of higher risk, symptom severity did *not* increase the likelihood that women would be incorrect in their assessments. This is an interesting point for practitioners to consider given that victims with mental health difficulties are often seen as less credible or reliable than are other victims (Finn and Stalans 1995, 1997). For this particular mental health issue common among survivors of IPV, this perception does not appear to hold water.

Also interesting is that for both forms of accuracy, women with more severe stalking histories were significantly more likely to be true positives than in other accuracy categories. Underscoring the relative strength of these findings, stalking accounted for a significant amount of variance in discriminating women who were correct from those who were incorrect even after accounting for physical and psychological abuse. That is, experiences of stalking contributed unique information to the determination of accuracy category membership over and above the information provided by forms of abuse more directly connected to the type of abuse victims were being asked to predict (psychological abuse in this study; physical abuse in the previous study). This finding seems particularly policy relevant, given that stalking in the absence of physical assault is often taken less seriously by practitioners, particularly in terms of decisions related to arrest and prosecution (Brewster 2001).

In conclusion, given the significant emotional and physical consequences of psychological and other forms of abuse, it is imperative that we learn more about how to predict and prevent them. In this quest, one sensible approach would be to explore the extent to which a variety of sources of information might potentially contribute to more accurate assessments of risk. To this end, while we emphasize the preliminary nature of our findings, we note that they fail to support concerns that victims are on average likely to distort the level of danger they face from their abusive partners. Certainly, not all victims will be accurate in their prediction of reabuse but then again, neither will all practitioners—existing research in fact suggests the two groups have comparable predictive abilities in this regard (Bennett Cattaneo 2007). With this in mind, we suggest that practitioners would do well to view their relationship with victims as a partnership to which each brings certain expertise, skills, and blind spots (Davies et al. 1998). Ultimately, this type of collaboration has the potential to enrich everyone's knowledge base—and thereby bring everyone's goal of ending the abuse within closer reach.

Acknowledgements We are grateful for the contributions made by everyone who has worked on this longitudinal project but particularly wish to thank Jane Murphy and Dorothy Lennig, and to express appreciation for the funding provided by The National Institute of Justice.

References

Arias, I. (1999). Women's response to physical and psychological abuse. In X. B. Arriagaa & S. Oskamp (Eds.), *Violence in intimate relationships* (pp. 139–161). Thousand Oaks, CA: Sage.

Arias, I., & Pape, K. T. (1999). Psychological abuse: Implications for adjustment and commitment to leave violent partners. *Violence & Victims, 14*(1), 55–67.

Bennett Cattaneo, L. (2007). Contributors to assessments of risk in intimate partner violence: How victims and professionals differ. *Journal of Community Psychology, 35*(1), 57–75.

Bennett Cattaneo, L., Bell, M. E., Goodman, L. A., & Dutton, M. A. (2007). Intimate partner violence victims' accuracy in assessing their risk of reabuse. *Journal of Family Violence* (in press).

Bennett Cattaneo, L., & Goodman, L. A. (2003). Victim-reported risk factors for continued abusive behavior: Assessing the dangerousness of arrested batterers. *Journal of Community Psychology, 31,* 349–369.

Bennett Cattaneo, L., & Goodman, L. A. (2007). New directions in IPV risk assessment: An empowerment approach to risk management. In K. Kendall-Tackett, & S. Giacomoni (Eds.), *Intimate partner violence* (in press).

Bennett, L., Goodman, L., & Dutton, M. A. (2000). Risk assessment among batterers arrested for domestic assault: The salience of psychological abuse. *Violence Against Women, 6*(11), 1190–1203.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD checklist. *Behaviour Research & Therapy, 34*, 669–673.

Brewster, M. P. (2001). Legal help-seeking experiences of former intimate-stalking victims. *Criminal Justice Policy Review, 12*(2), 91–112.

Bronfenbrenner, U. (1988). Interacting systems in human development: Research paradigms: Present and future. In N. Bolger, A. Caspi, G. Downey, & M. Moorehouse (Eds.), *Persons in context: Developmental processes. Human development in cultural and historical contexts* (pp. 25–49). New York: Cambridge University Press.

Brown, J. (1997). Working toward freedom from violence: The process of change in battered women. *Violence Against Women, 3*(1), 5–26.

Campbell, J. C. (1995). Prediction of homicide of and by battered women. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp. 96–113). Thousand Oaks, CA: Sage.

Cohen, S., Mermelstein, R., Kamarck, T., & Hoberman, H. M. (1985). Measuring the functional components of social support. In I. G. Sarason, & B. R. Sarason (Eds.), *Social support: Theory, research, and applications* (pp. 73–94). Boston: Martinus Nijhoff.

Davies, J., Lyon, E., & Monti-Catania, D. (1998). *Safety planning with battered women: Complex lives/difficult choices*. Thousand Oaks, CA: Sage Publications, Inc.

Finn, M. A., & Stalans, L. J. (1995). Police referrals to shelters and mental health treatment: Examining their decisions in domestic assault cases. *Crime & Delinquency, 41*(4), 467–480.

Finn, M. A., & Stalans, L. J. (1997). The influence of gender and mental state on police decisions in domestic assault cases. *Criminal Justice and Behavior, 24*(2), 157–176.

Follingstad, D. R., Rutledge, L. L., Berg, B. J., Hause, E. S., & Polek, D. S. (1990). The role of emotional abuse in physically abusive relationships. *Journal of Family Violence, 5*(2), 107–120.

Gondolf, E. W., & Heckert, D. A. (2003). Determinants of women's perceptions of risk in battering relationships. *Violence & Victims, 18,* 371–386.

Goodman, L. A., Dutton, M. A., Weinfurt, K., & Cook, S. (2003). The Intimate Partner Violence Strategies Index. *Violence Against Women, 9*(2), 163–186.

Hall Smith, P., Thornton, G. E., DeVellis, R., Earp, J., & Coker, A. L. (2002). A population-based study of the prevalence and distinctiveness of battering, physical assault, and sexual assault in intimate relationships. *Violence Against Women, 8*(10), 1208–1232.

Heckert, D. A., & Gondolf, E. W. (2004). Battered women's perceptions of risk versus risk factors and instruments in predicting repeat reassault. *Journal of Interpersonal Violence, 19,* 778–800.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.

House, J. S., & Kahn, R. L. (1985). Measures and concepts of social support. In S. Cohen, & S. L. Syme (Eds.), *Social support and health* (pp. 83–108). New York: Academic.

Jacobson, N. S., Gottman, J. M., Gortner, E., Berns, S., & Shortt, J. W. (1996). Psychological factors in the longitudinal course of battering: When do the couples split up? When does the abuse decrease? *Violence & Victims, 11*(4), 371–392.

Katz, J., & Arias, I. (1999). Psychological abuse and depressive symptoms in dating women: Do different types of abuse have differential effects? *Journal of Family Violence, 14*(3), 281–295.

Kelly, V. A. (2004). Psychological abuse of women: A review of the literature. *The Family Journal, 12*(4), 383–388.

Langford, D. R. (1996). Predicting unpredictability: A model of women's processes of predicting battering men's violence. *Scholarly Inquiry for Nursing Practice, 10,* 371–385.

Litz, B. T., & Keane, T. M. (1989). Information processing in anxiety disorders: Application to the understanding of posttraumatic stress disorder. *Clinical Psychology Review, 9,* 243–257.

Marshall, L. L. (1999). Effects of men's subtle and overt psychological abuse on low-income women. *Violence & Victims, 14*(1), 69–88.

Murphy, C. M., & Hoover, S. A. (1999). Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence & Victims, 14,* 39–53.

O'Leary, K. D. (1999). Psychological abuse: A variable deserving critical attention in domestic violence. *Violence & Victims, 14*(1), 3–23.

Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model*. New York: Springer.

Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research, Practice, Training, 19,* 276–288.

Roehl, J., O'Sullivan, C., Webster, D., & Campbell, J. (2005). *Intimate partner violence risk assessment validation study: The RAVE study practitioner summary and recommendations*. Washington, DC: National Institute of Justice.

Stith, S. M., Smith, D. B., Penn, C. E., Ward, D. B., & Tritt, D. (2004). Intimate partner physical abuse perpetration and victimization risk factors: A meta-analytic review. *Aggression & Violent Behavior, 10*(1), 65–98.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scale (CTS2): Development and preliminary psychometric data. *Journal of Family Issues, 17*(3), 283–316.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Tjaden, P., & Thoennes, N. (2000). *Extent, nature, and consequences of intimate partner violence: Findings from the National Violence Against Women Survey*. Washington, DC: National Institute of Justice.

Tolman, R. M. (1989). The development of a measure of a psychological maltreatment of women by their male partners. *Violence & Victims, 4*(3), 159–177.

Tolman, R. M. (1999). The validation of the Psychological Maltreatment of Women Inventory. *Violence & Victims, 14*(1), 25–37.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist: Reliability, validity, and diagnostic utility*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weisz, A. N., Tolman, R. M., & Saunders, D. G. (2000). Assessing the risk of severe violence: The importance of survivors' predictions. *Journal of Interpersonal Violence, 15,* 75–90.